Literal Homomorphisms of OL-languages

by

K. Culik II and J. Opatrný

Research Report CS-73-25
September 1973

Department of Applied Analysis and
Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Faculty of Mathematics

University of Waterloo

Waterloo, Ontario

Canada

Department of Applied Analysis

&

Computer Science

Literal Homomorphisms of OL-languages

by
K. Culik II and J. Opatrný

Research Report CS-73-25
September 1973

Department of Applied Analysis and
Computer Science
University of Waterloo
Waterloo, Ontario, Canada

This work was supported by the National Research Council of Canada Grant No. A7403.

ABSTRACT

Impossibility to distinguish all different states of the cells of a developing organism during a biological experiment motivates the study of literal homomorphic images (total codings) of languages generated by OL-systems and their subclasses, e.g. propagating or deterministic OL-systems. In this paper the properties of corresponding families of languages are studied and the relations among them and also their relation to the families of context-free and regular languages are investigated.

1. Introduction Lindenmayer systems or L-systems were introduced in [7] as a tool for modelling the development of filamentous organisms.

An L-system describes the development of a filamentous organism which is represented by a string of symbols, each symbol representing a state of a cell. Possible changes of states of cells, as well as their divisions are described by a finite set of rewriting rules. Rules have to be applied simultaneously to all symbols in the string to obtain a new stage of the development of an organism.

OL-systems are models for the development without any cell interaction. Properties of OL-systems and OL-languages has been studied in [8], [9].

When observing a cell in an actual experiment, usually not all of its properties can be determined and two cells which look alike for an observer may actually behave differently. Mathematically speaking, the observation of a string of cells, under the assumption that each individual cell (its state) is represented by a symbol gives us only a literal homomorphic image (ℓh-image) of the string. The above motivated the investigation of ℓh-images (also called total codings in [11]) of OL-languages.

In this paper we are studying the descriptive power of ℓh-images of OL-languages and their various subclasses and their relations to well known families of languages. For example, two problems which were open for some time have been solved. It is shown that every CF language which contains a nonempty string is ℓh-image of an OL-language, and that the family of the ε-free regular languages is properly contained in the family of the ℓh-images of languages generated by propagating

OL-systems with finite number of starting words (i.e. generated by PFOL-systems).

To prove proper inclusion of certain families of languages we are introducing a new concept of essentially deterministic languages. We hope that this concept may help to attack rather difficult questions of membership for similar types of languages.

In section 5 we briefly mention some closure properties of ℓh-images of OL-languages.

2. Preliminaries We shall assume that the reader is familiar with basic notation and notions of formal language theory, e.g. [6].

Now, we review the definition of OL-systems and OL-languages and introduce some notation used throughout the paper.

Notation The length of a string α is denoted by $|\alpha|$. The empty string is denoted by ε.

Definition 1 An OL-system G is an ordered 3-tuple (Σ,P,σ) where:

(i) Σ is a finite nonempty set of symbols called the alphabet.

(ii) P is the finite set of ordered pairs from $\Sigma \times \Sigma^*$ called the productions. A production (a,α) where $a \in \Sigma$, $\alpha \in \Sigma^*$ is usually written as $a \to \alpha$.

(iii) $\sigma \in \Sigma^*$ is the initial string.

Any OL-system has to be complete which means that for every $a \in \Sigma$ there must exist a string $\alpha \in \Sigma^*$ such that $(a,\alpha) \in P$.

Given an OL-system $G = (\Sigma, P, \delta)$, let $\alpha = a_1 a_2 \ldots a_n$, $a_i \in \Sigma$ and $\beta \in \Sigma^*$. α is said to directly derive β (in an OL-system G) written $\alpha \underset{G}{\Longrightarrow} \beta$ if there exist $\beta_1, \beta_2, \ldots, \beta_n \in \Sigma^*$ such that $\beta = \beta_1 \beta_2 \ldots \beta_n$ and $a_1 \to \beta_1$, $a_2 \to \beta_2, \ldots, a_n \to \beta_n$ are productions in P.

Let relation $\overset{k}{\underset{G}{\Longrightarrow}}$ for $k \geq 0$ be defined on Σ^* as follows:

(i) $\alpha \overset{0}{\underset{G}{\Longrightarrow}} \beta$ iff $\alpha = \beta$;

(ii) $\alpha \overset{k}{\underset{G}{\Longrightarrow}} \beta$ for $k \geq 1$ iff there exists $\gamma \in \Sigma^*$ such that $\alpha \underset{G}{\Longrightarrow} \gamma$ and $\gamma \overset{k-1}{\underset{G}{\Longrightarrow}} \beta$.

Let $\underset{G}{\Longrightarrow}^*$ be the reflexive and transitive closure of the relation $\underset{G}{\Longrightarrow}$ on Σ^*, i.e. $\alpha \underset{G}{\Longrightarrow}^* \beta$ iff there exists $k \geq 0$ so that $\alpha \overset{k}{\underset{G}{\Longrightarrow}} \beta$. We say that $\alpha \Longrightarrow^+ \beta$ iff $\alpha \overset{k}{\Longrightarrow} \beta$ for some $k \geq 1$. Language L generated by an OL-system G is denoted L(G) and is defined to be the set $\{\alpha \in \Sigma^* : \sigma \underset{G}{\Longrightarrow}^* \alpha\}$.

<u>Definition 2</u> An FOL-system G is an ordered 3-tuple (Σ, P, F) where:

(i) Σ and P are the same as in the definition of OL-system.

(ii) Completeness and relations $\underset{G}{\Longrightarrow}$, $\overset{k}{\underset{G}{\Longrightarrow}}$, and $\overset{*}{\underset{G}{\Longrightarrow}}$ have the same meaning as in the definition of OL-systems. FOL-language L(G) is defined to be the set $\{\alpha : \sigma \overset{*}{\underset{G}{\Longrightarrow}} \alpha, \sigma \in F\}$.

<u>Definition 3</u> Let $G = (\Sigma, R, F)$ be an OL-system or an FOL-system. G is called <u>deterministic</u> (abbreviated D) if for every $a \in \Sigma$ there exists exactly one $\alpha \in \Sigma^*$ such that $(a, \alpha) \in R$.

<u>Definition 4</u> Let $G = (\Sigma, R, F)$ be an OL-system or an FOL-system. G is called <u>propagating</u> (abbreviated P) if $R \subseteq \Sigma \times \Sigma^+$.

Notation If an OL or an FOL-system has any or several properties defined above, we compound their abbreviations together with OL or FOL, respectively. E.g. G is a propagating deterministic FOL-system is abbreviated by G is a PDFOL-system.

Definition 5 Let Σ, Γ be finite sets of symbols. Literal homomorphism from Σ^* to Γ^*, abbreviated ℓ-homomorphism is a total function $h: \Sigma \to \Gamma$ extended to strings from Σ^* by:

(i) $h(\varepsilon) = \varepsilon$;

(ii) $h(\alpha\ \beta) = h(\alpha)h(\beta)$ for $\alpha, \beta \in \Sigma^*$.

Note In [11] an ℓ-homomorphism is called a total coding.

Definition 6 If L is a language over Σ and t is an ℓ-homomorphism from Σ to Γ then t(L) is the set $\{t(\alpha) : \alpha \in L\}$.

Notation The family of languages of type X is denoted by $\mathcal{L}_X$, e.g. $\mathcal{L}_{CF}$ is the family of CF (context free) languages.

We say that sets A, B are incomparable if $A \nsubseteq B$ and $B \nsubseteq A$.

Definition 7 Let $\mathcal{L}_X$ be a family of languages. The family of literal homomorphic images (abbreviated ℓh-images) of $\mathcal{L}_X$, denoted $\mathcal{L}_{hX}$ is the set $\{t(L) : L \in \mathcal{L}_X$ and t is an ℓ-homomorphism defined for all words of L$\}$.

3. The ℓh-images of OL-languages and FOL-languages

Theorem 1 $\mathcal{L}_{hOL} = \mathcal{L}_{hFOL}$.

Proof Let $L = t(L(G))$ where $G = (\Sigma, P, F)$ is an FOL-system and t is an ℓ-homomorphism, $t: \Sigma^* \to \Gamma^*$. Let $\sigma \in F$, $\sigma = a_1 a_2 \ldots a_n$ where $a_i \in \Sigma$ for $i = 1,2,\ldots,n$. Let $a_1', a_2', \ldots, a_n'$ be pairwise different symbols not in Σ.

Construct OL-system $G' = (\Sigma',P',\sigma')$ where $\Sigma' = \Sigma \cup \{a_1',a_2',\ldots,a_n'\}$, $\sigma' = a_1'a_2'\ldots a_n'$ and $P' = P \cup \{a_1' \to \beta\colon \beta \in F\} \cup \{a_j' \to \varepsilon\colon j = 2,3,\ldots,n\}$. Define ℓ-homomorphism $t'\colon(\Sigma')^* \to \Gamma^*$ as follows:

(i) $t'(a) = t(a)$ for all $a \in \Sigma$;

(ii) $t'(a_j') = t(a_j)$ for $j = 1,2,\ldots,n$.

Since system G is complete and we have constructed a production for each symbol in $\Sigma' - \Sigma$, also G' is complete. From the construction of P' follows that $\alpha \underset{G}{\Rightarrow}^* \beta$ where $\alpha \in F$ if and only if $\sigma' \underset{G'}{\Rightarrow}^+ \beta$. Furthermore, $t'(\beta) = t(\beta)$ and $t'(\sigma') \in F$. So $t(L(G)) = t'(L(G'))$. □

<u>Theorem 2</u> $\mathcal{L}_{OL} \subsetneq \mathcal{L}_{hOL}$.

<u>Proof</u> Inclusion of $\mathcal{L}_{OL}$ in $\mathcal{L}_{hOL}$ is trivial. By theorem 1 the inclusion is proper since $\mathcal{L}_{OL}$ is not closed under union (see [9]), and the closure of $\mathcal{L}_{hFOL}$ under union is obvious.

<u>Theorem 3</u> $\mathcal{L}_{CF} - \{\varepsilon\} \subsetneq \mathcal{L}_{hOL}$.

<u>Proof</u> Let L be a CF language, $L \neq \{\varepsilon\}$. Let $G = (N,T,P,S)$ be a reduced ε-free CF grammar generating L (we allow the production $S \to \varepsilon$ if S does not occur on the right-side of any production). Construct OL-system $G' = (\Sigma,P',\sigma)$ in the following way. Let $\Sigma = T \times (N \cup \{\varepsilon\})$. Let f be the homomorphism on $(N \cup T)^*$ defined by:

(i) $f(a) = a$ for $a \in T$;

(ii) $f(A) = (a_1,A)(a_2,\varepsilon)\ldots(a_n,\varepsilon)$ where $a_1a_2\ldots a_n$ is any fixed string in T^+ such that $A \Rightarrow^* a_1a_2\ldots a_n$, $a_1,a_2,\ldots,a_n \in T$. Such a string obviously exists for every $A \in N$ because of the assumption above.

Let $\sigma = f(S)$. Finally construct P' :

(i) if $A \to \alpha \in P$ then $(a,A) \to f(\alpha) \in P'$ for every $a \in T$;

(ii) $(a,\varepsilon) \to \varepsilon \in P$ for every $a \in T$;

(iii) $a \to a \in P$ for every $a \in T$.

Obviously, G' is complete and thus it is an OL-system. Let t be the homomorphism defined by:

(i) $t(a) = a$ for $a \in T$;

(ii) $t(a,X) = a$ for every $a \in T$ and $X \in N \cup \{\varepsilon\}$.

Clearly t is ℓ-homomorphism and $t(L(G)) = L(G)$.

It has been shown in [9] that $\mathcal{L}_{OL}$ and $\mathcal{L}_{CF}$ are incomparable. Thus by Theorem 2 the inclusion of $\mathcal{L}_{CF}$ in $\mathcal{L}_{hOL}$ is proper. □

Let the family of index languages [1] be denoted by $\mathcal{L}_{INDEX}$ and let the family of context-sensitive languages be denoted by $\mathcal{L}_{CS}$. Using results from [2] we have the following theorem.

Theorem 4 $\mathcal{L}_{hOL} \subsetneqq \mathcal{L}_{INDEX} \subsetneqq \mathcal{L}_{CS}$.

Proof The family of FMOL-languages defined in [2] is equal to the closure of $\mathcal{L}_{OL}$ under finite substitution. Literal homomorphism is a special case of finite substitution and therefore $\mathcal{L}_{hOL} \subseteq \mathcal{L}_{FMOL}$. It has been also shown in [2] that $\mathcal{L}_{FMOL} \subsetneqq \mathcal{L}_{INDEX}$ and by [1] $\mathcal{L}_{INDEX} \subsetneqq \mathcal{L}_{CS}$. □

4. <u>ℓh-images of propagating OL-languages</u> In the proofs of Theorems 1 and 3 we used essentially erasing productions (i.e. productions of the form $a \to \varepsilon$). Now, we will have to use a different techniques to prove some results on literal-homomorphic images of POL-languages and PFOL-languages.

Let the set of all ε-free regular sets be denoted by $\mathcal{L}_{\mathrm{REGULAR}}$.

<u>Theorem 5</u> $\mathcal{L}_{\mathrm{REGULAR}} \subsetneq \mathcal{L}_{\mathrm{hPFOL}}$.

<u>Proof</u> First, we will describe the proof informally. Let M be a finite automaton accepting L in $\mathcal{L}_{\mathrm{REGULAR}}$. A finite automaton can be represented by a state diagram in the usual manner as in [6]. We will construct a PFOL-system generating encodings of all possible paths in the state diagram leading from the starting state to a final state. Each path is encoded by a string of pairs such that the first components of the pairs are the labels of the edges on the path and the second components are the nodes of the path except of the first one (starting state). The set of starting strings of the constructed PFOL-system will be the set of encodings of all shortest paths (i.e. without any loop) from the starting state to a final state. Productions of the PFOL- system will allow to add the encoding of any loop wherever it is possible. $\mathcal{L}$-homomorphism t will map a pair onto its first component and thus will map an encoding of a path generated by the PFOL-system to a string of L.

Now we proceed with the formal construction. Let L be an ε-free regular language. Let $M = (K,\Sigma,\delta,q_0,F)$ be a finite automaton accepting L (see [6], p.26). We can suppose without loss of generality that $\delta(q,\varepsilon) = \phi$ for all $q \in K$, and $q_0 \notin \delta(q,a)$ for any $a \in \Sigma$, $q \in K$. Let R be the set of all strings α satisfying the following conditions:

(i) $\alpha = (a_1,p_1)(a_2,p_2)\ldots(a_n,p_n)$ where $n \geq 1$, $a_i \in \Sigma$, $p_i \in K$ for $i = 1,2,\ldots,n$;

(ii) $p_1 \in (q_0,a_1)$ and $p_i \in (p_{i-1},a_i)$ for $i = 2,3,\ldots,n$;

(iii) if $i \neq j$ then $p_i \neq p_j$ for $i,j = 1,2,\ldots,n$;

(iv) $p_n \in F$.

Let P_q be the set of all strings β satisfying the following conditions:

(i) $\beta = (b_1,r_1)(b_2,r_2)\ldots(b_m,r_m)$ where $m \geq 1$, $b_i \in \Sigma$, $r_i \in K$ for $i = 1,2,\ldots,m$;

(ii) $r_1 \in \delta(q,b_1)$, $r_i \in \delta(r_{i-1},b_i)$ for $i = 2,3,\ldots,m$;

(iii) if $i \neq j$ then $r_i \neq r_j$ for $i,j = 1,2,\ldots,m$;

(iv) $r_m = q$.

Clearly, R and P_q are finite for any $q \in K$. Define ℓ-homomorphism $t:(\Sigma \times K)^* \to \Sigma^*$ by $t((a,q)) = a$ for every $a \in \Sigma$, $q \in K$. Let G be an FPOL-system, $G = (\Sigma \times K,P,R)$, where

$$P = \{(a,q) \to (a,q)\alpha : a \in \Sigma,\ q \in K,\ P_q \neq \phi,\ \ \alpha \in P_q\} \cup \{(a,q) \to (a,q) : a \in \Sigma,\ q \in K\}.$$

From the construction of G follows that if $\alpha \overset{*}{\underset{G}{\Rightarrow}} \beta$ where $\alpha \in R$ then $(q_0,t(\beta)) \mathrel{|\!\frac{*}{M}} (q,\varepsilon)$, $q \in F$ and if $(q_0,\gamma) \mathrel{|\!\frac{*}{M}} (q,\varepsilon)$ where $q \in F$ then there exist $\alpha \in R$ and β such that $\alpha \overset{*}{\underset{G}{\Rightarrow}} \beta$ and $t(\beta) = \gamma$. $\square$

Now, we will study properties of ℓh-images of propagating deterministic OL-languages. It has been shown in [9] that OL-languages do not include all ε-free finite sets. However, we will show that all ε-free finite sets, denoted by $\mathscr{L}_{\text{FINITE}}$, are included in ℓh-images of the simplest type of OL-languages, namely in ℓh-images of PDOL-languages.

<u>Theorem 6</u> $\mathscr{L}_{FINITE} \subsetneq \mathscr{L}_{hPDOL}$.

<u>Proof</u> Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a finite subset of Σ^+ where Σ is a finite alphabet. Let $\alpha_i = a_{i1}a_{i2}\ldots a_{is_i}$, where $s_i \geq 1$ for $1 \leq i \leq n$, and $a_{ij} \in \Sigma$ for $1 \leq i \leq n$, $1 \leq j \leq s_i$. We may suppose without loss of generality that $s_i \leq s_{i+1}$ for $1 \leq i \leq n-1$.

Let $\Delta = \{b_{ij} : 1 \leq i \leq n, 1 \leq j \leq s_i\}$ be a set of parwise distinct symbols. We construct OL-system $G = (\Delta, P, \sigma)$ where

$$\sigma = b_{11}b_{12}\ldots b_{1s_1} \text{ and}$$

$$P = \{b_{ij} \to b_{i+1j} : 1 \leq i \leq n, 1 \leq j \leq s_i - 1\} \cup$$

$$\cup \{b_{is_i} \to b_{i+1s_i} b_{i+1s_i+1} \ldots b_{i+1s_{i+1}} : 1 \leq i \leq n-1\}$$

$$\cup \{b_{nj} \to b_{nj} : 1 \leq j \leq s_n\}.$$

Clearly, G is PDOL-system. Now, we define ℓ-homomorphism t, $t : \Delta^* \to \Sigma^*$, $t(b_{ij}) = a_{ij}$ for $1 \leq i \leq n$, $1 \leq j \leq s_i$. Then, obviously $t(L(G)) = A$. □

<u>Notation</u> The number of elements of a finite set A is denoted by $||A||$.

For a language L, the subset of L $\{\alpha \in L : |\alpha| = n\}$ is denoted by $L^{(n)}$.

<u>Lemma 1</u> Let L be an hPDFOL-language. Then there exists a constant c (independent on n) such that for every n

$$||L^{(n)}|| \leq c.$$

<u>Proof</u> Since any hPDFOL-language is the union of a finite number of hPDOL-languages, it is enough to prove Lemma 1 for hPDOL-languages.

Let $L = t(L(G))$ where $G = (\Sigma, P, \sigma)$ be a PDOL-system with $||\Sigma|| = m$ and t be an ℓ-homomorphism.

Case 1. Suppose there exist $\alpha_1, \alpha_2, \ldots, \alpha_{m+1} \in L(G)$ such that $\alpha_j \Rightarrow \alpha_{j+1}$ and $|\alpha_j| = |\alpha_{j+1}|$ for $j = 1,2,\ldots,m$. Since G is a deterministic system with only m different symbols in its alphabet, we have $|\alpha_1| = |\beta|$ for every β such that $\alpha_1 \Rightarrow^* \beta$ and since t is length preserving homomorphism, $||L^{(n)}|| = 0$ for $n > |\alpha_1|$. In this case $c = \max\{L^{(j)} : 1 \le j \le |\alpha_1|\} + 1$.

Case 2. Suppose that for any $\alpha_1, \alpha_2, \ldots, \alpha_{m+1}$ in $L(G)$, $\alpha_j \Rightarrow \alpha_{j+1}$ for $j = 1,2,\ldots,m$ implies $|\alpha_{m+1}| > |\alpha_1|$.

Then, clearly, $||L^{(n)}|| < (m+2)$ for every $n \ge 1$. □

Lemma 2 $\mathcal{L}_{REGULAR}$ is incomparable to $\mathcal{L}_{hPDOL}$ and to $\mathcal{L}_{hPDFOL}$. Also, $\mathcal{L}_{CF}$ is incomparable to $\mathcal{L}_{hPDOL}$ and to $\mathcal{L}_{hPDFOL}$.

Proof Let $G = (\{a\}, \{a \to aa\}, a)$, and let t be the identity. G is an PDOL-system, t is an ℓ-homomorphism, and $t(L(G)) = \{a^{2^n} : n \ge 0\}$ which is neither regular nor context-free language.

Let $L_1 = a^+b^+$. Clearly, $||L_1^{(n)}|| = n-1$ and therefore by Lemma 1, L_1 is neither in $\mathcal{L}_{hPDOL}$ nor in $\mathcal{L}_{hPDFOL}$. □

Theorem 7 $\mathcal{L}_{hPDFOL} \subsetneq \mathcal{L}_{hPFOL}$.

Proof The inclusion is trivial. It is proper by Theorem 5 and Lemma 2. □

Theorem 8 $\mathcal{L}_{hPDOL} \subsetneq \mathcal{L}_{hPOL}$.

Proof Inclusion of $\mathcal{L}_{hPDOL}$ in $\mathcal{L}_{hPOL}$ is trivial. In the proof of the Lemma 2 we have shown that the language $L_1 = a^+b^+$ is not in $\mathcal{L}_{hPDOL}$. However,

L_1 is generated by POL-system $G = (\{a,b\},\{a \to aa,\ a \to a,\ b \to bb,\ b \to b,\ ab\}$; thus L_1 is in $\mathcal{L}_{POL}$ and also in $\mathcal{L}_{hPOL}$. □

Now, we will show the proper inclusions of $\mathcal{L}_{hPDOL}$ in $\mathcal{L}_{hPDFOL}$ and of $\mathcal{L}_{hPOL}$ in $\mathcal{L}_{hPFOL}$. To do so we will first define and study properties of "essentially nondeterministic" hPOL-systems and "essentially deterministic" hPOL-languages. Informally, an OL-system G is essentially nondeterministic with respect to ℓ-homomorphism t, if there is a symbol in alphabet of G which occurs in infinitely many words of L(G) and from which we can derive in k steps, for some $k \geq 1$, two strings α, β such that $t(\alpha) \neq t(\beta)$.

Language L is an essentially deterministic hPOL-language if there is no POL-system G and ℓ-homomorphism t such that G is essentially nondeterministic with respect to t and $L = t(L(G))$.

After showing some properties of essentially deterministic languages we will exhibit two essentially deterministic hPDOL-languages whose union is neither in $\mathcal{L}_{hPDOL}$ nor in $\mathcal{L}_{hPOL}$.

Definition 8 POL-system $G = (\Sigma,P,\sigma)$ is an essentially nondeterministic hPOL-system with respect to ℓ-homomorphism t if there exists $a \in \Sigma$ and $\gamma_1,\gamma_2 \in \Sigma^*$ so that a occurs in infinitely many words of L(G), $a \overset{k}{\underset{G}{\Rightarrow}} \gamma_1$, $a \overset{k}{\underset{G}{\Rightarrow}} \gamma_2$ for some $k \geq 1$ and $t(\gamma_1) \neq t(\gamma_2)$.

Definition 9 Language L is an essentially deterministic hPOL-language if for every pair of POL-system G and ℓ-homomorphism t, such that $L = t(L(G))$, G is not an essentially nondeterministic POL-system with respect to t.

Let L be an essentially deterministic h-POL language, $L = h(L(G))$ for POL-system G and homomorphism h. We will show in Lemma 4 that L can be expressed as the union of finite number of hPDOL-languages, each of them generated by a system preserving the structure of derivations in G. We will need this in the proof of Lemma 9. We first need the following auxiliary result.

<u>Lemma 3</u> Let G be an POL-system, $G = (\Sigma,P,\sigma)$. For any $c > 0$ there exists $j_c > 0$ such that for any $\alpha \in L(G)$ either $\sigma \overset{i}{\underset{G}{\Rightarrow}} \alpha$ with $i < j_c$ or there exists β such that $|\beta| \geq c$ and $\sigma \overset{j_c}{\underset{G}{\Rightarrow}} \beta \underset{G}{\Rightarrow}^* \alpha$.

<u>Proof</u> Let $n = ||\Sigma||$ and let $j_c = n^{c+1}$. Suppose that $\alpha \in L(G)$ cannot be derived within j_c steps from σ and that $\sigma \overset{j_c}{\Rightarrow} \beta \Rightarrow^* \alpha$ where $|\beta| < c$. System G is propagating, therefore, for any string γ, such that $\sigma \Rightarrow^* \gamma \Rightarrow^* \beta$, we have $|\gamma| < c$. There are less than n^{c+1} different strings of the length smaller than c over Σ and therefore there exist $i_1 \geq 0$, $i_2 > 0$ and $\delta \in \Sigma^*$ such that $\sigma \overset{i_1}{\Rightarrow} \delta \overset{i_2}{\Rightarrow} \delta \xRightarrow{j_c-i_1-i_2} \beta \Rightarrow^* \delta$. Thus also $\sigma \overset{i_1}{\Rightarrow} \delta \xRightarrow{j_c-i_1-i_2} \beta \Rightarrow^* \alpha$.

Let β_1 be the string $\sigma \xRightarrow{j_c-i_2} \beta \overset{i_2}{\Rightarrow} \beta_1 \Rightarrow^* \alpha$. If $|\beta_1| < c$ then we can repeat the process above until we obtain β' such that $\sigma \overset{j_c}{\Rightarrow} \beta' \Rightarrow^* \alpha$ and $|\beta'| \geq c$. □

<u>Lemma 4</u> Let L be an essentially deterministic hPOL-language. Let $L = t(L(G))$ where G is a POL-system and t is an ℓ-homomorphism. Then there exist $n \geq 1$, ℓ-homomorphismus t' and PDOL-systems $G_1, G_2, \ldots, G_n$ such that $L = \bigcup_{i=1}^{n} t(t'(L(G_i)))$ and if $\alpha \overset{j}{\underset{G_i}{\Rightarrow}} \beta$ for some $j \geq 1$, $1 \leq i \leq n$, $\alpha \in L(G_i)$ then $t'(\alpha) \overset{j}{\underset{G}{\Rightarrow}} t'(\beta)$.

Proof Let $G = (\Sigma,P,\sigma)$. We may suppose that every symbol of Σ occurs in some string of $L(G)$.

Suppose that for a in Σ there exist two productions $a \to \alpha$, $a \to \beta$ in P, $\alpha \neq \beta$. Suppose, also, that a occur in infinitely many words of $L(G)$. Since L is an essentially deterministic hPOL-language, for any $k \geq 0$ and $\gamma_1, \gamma_2 \in \Sigma^+$ if $a \Rightarrow \alpha \overset{k}{\underset{G}{\Rightarrow}} \gamma_2$ and $a \underset{G}{\Rightarrow} \beta \overset{k}{\underset{G}{\Rightarrow}} \gamma_2$ then $t(\gamma_1) = t(\gamma_2)$. Hence, system $G' = (\Sigma, P-\{(a,\alpha)\},\sigma)$ also generates language $L(G)$.

By finite number of repetitions of the deletion of one of the pair of the considered productions we can obtain system $G_1 = (\Sigma,P_1,\sigma)$ and Σ_1,Σ_2 such that:

(i) $\Sigma_1 \cap \Sigma_2 = \phi$, $\Sigma = \Sigma_1 \cup \Sigma_2$.

(ii) $L(G_1) = L(G)$.

(iii) There is exactly one production in P_1 for any symbol of Σ_2.

(iv) Symbols of Σ_1 occur only in finite number of words of $L(G)$.

(v) $\alpha \overset{k}{\underset{G_1}{\Rightarrow}} \beta$ iff $\alpha \overset{k}{\underset{G}{\Rightarrow}} \beta$.

Let $A = \{\alpha : \alpha \in L(G) - \Sigma_2^+\}$. By (iv) A is finite. Let $c = \max\{|\alpha| : \alpha \in A\} + 1$. By Lemma 3 there exists constant j_c such that any $\alpha \in L(G_1)$ either can be derived from σ within j_c steps, or there exists β such that $|\beta| \geq c$ and $\sigma \overset{j_c}{\underset{G_1}{\Rightarrow}} \beta \underset{G_1}{\Rightarrow}^* \alpha$. Therefore, $A \subseteq \{\alpha : \sigma \overset{i}{\underset{G_1}{\Rightarrow}} \alpha,\ 0 \leq i \leq j_c\}$.

Let $B = \{\alpha : |\alpha| \geq c$ and $\sigma \overset{j_c}{\underset{G_1}{\Rightarrow}} \alpha\}$. Let $C = \{\alpha : \sigma \overset{j_c}{\underset{G_1}{\Rightarrow}} \alpha\}$. Clearly, B and C are finite, $B \subseteq C$. Let $G_\alpha = (\Sigma,P_1,\alpha)$ for every $\alpha \in L(G_1)$. We have $L(G_1) = \bigcup_{\alpha\in B} L(G_\alpha) \cup \{\alpha : \sigma \overset{i}{\underset{G_1}{\Rightarrow}} \alpha,\ 0 \leq i \leq j_c\}$. Let P' be any fixed subset of P_1 such that there is exactly one production in P' for every symbol of Σ.

By (iii) $P' \cap (\Sigma_2 \times \Sigma_2^+) = P_1 \cap (\Sigma_2 \times \Sigma_2^+)$. Let $G'_\alpha = (\Sigma, P', \alpha)$ for every $\alpha \in L(G_1)$. Clearly, G'_α is a PDOL-system for every $\alpha \in L(G_1)$. Since G_1 is a propagating OL-system $L(G_\alpha) \subseteq \Sigma_2^+$ for every $\alpha \in B$. Hence, $L(G_\alpha) = L(G'_\alpha)$ for every $\alpha \in B$ and $L(G_1) = \bigcup_{\alpha \in B} L(G'_\alpha) \cup \{\alpha : \sigma \overset{i}{\underset{G_1}{\Rightarrow}} \alpha,\ 0 \le i \le j_c\} = \bigcup_{\alpha \in C} L(G'_\alpha) \cup \{\alpha : \sigma \overset{i}{\underset{G_1}{\Rightarrow}} \alpha,\ 0 \le i \le j_c\}$, and if $\alpha \overset{k}{\underset{G_\alpha}{\Rightarrow}} \beta$ then $\alpha \overset{k}{\underset{G_1}{\Rightarrow}} \beta$.

Let $\alpha \in C$, and let $(\alpha_1^1, \alpha_2^1, \ldots, \alpha_{j_c}^1), \ldots, (\alpha_1^{m_\alpha}, \alpha_2^{m_\alpha}, \ldots, \alpha_{j_c}^{m_\alpha})$, where $m_\alpha \ge 1$, be all sequences of strings of $L(G_1)$ such that:

(i) $\alpha_j^i \underset{G}{\Rightarrow} \alpha_{j+1}^i$ for $1 \le i \le m_\alpha$, $1 \le j \le j_c - 1$;

(ii) $\alpha_1^i = \sigma$, $\alpha_{j_c}^i = \alpha$ for $1 \le i \le m_\alpha$.

For every $\alpha \in C$ and all i, $1 \le i \le m_\alpha$, we can easily construct a DPOL-system $G_\alpha^i = (\Gamma_\alpha^i, P_\alpha^i, \omega_\alpha^i)$ and ℓ-homomorphism t' such that :

(i) If $\omega_\alpha^i \overset{j}{\underset{G_\alpha^i}{\Rightarrow}} \beta$ for $0 \le j \le j_c - 1$ then $t'(\beta) = \alpha_{j+1}^i$.

(ii) For any $\alpha, \beta \in L(G'_\alpha)$, $\alpha \underset{G'_\alpha}{\Rightarrow} \beta$ iff $\alpha \underset{G_\alpha^i}{\Rightarrow} \beta$.

Thus, if $\alpha \overset{k}{\underset{G_\alpha^i}{\Rightarrow}} \beta$ for some $i, k \ge 1$, $\alpha \in L(G_i^\alpha)$, then $t'(\alpha) \overset{k}{\underset{G_1}{\Rightarrow}} t'(\beta)$.

Let $n = \sum_{\alpha \in C} m_\alpha$. Then $L(G) = L(G_1) = \bigcup_{\alpha \in C} (\bigcup_{i=1}^{m_\alpha} t'(L(G_\alpha^i)))$, and if $\alpha \overset{k}{\underset{G_\alpha^i}{\Rightarrow}} \beta$ for some $i, k \ge 1$, $\alpha \in C$, then $t(\alpha) \overset{k}{\underset{G}{\Rightarrow}} t'(\beta)$. □

We have shown that any essentially deterministic hPOL-language can be expressed as the union of a finite number of deterministic languages.

However, there are essentially deterministic languages which are not deterministic, i.e. for which $n > 1$ is necessary. $\{a^{3^n} : n \geq 1\} \cup \{ b^n c^n d^n : n \geq 0\} \cup \{a\}$ is such a language as will follow from Lemma 9. Also, there are deterministic languages which are not essentially deterministic, e.g. a^+ is such a language.

Notation For any language L, $\min\{|w| : w \in L\}$ is denoted by $\min(L)$.

Definition 10 Let $G = (\Sigma,P,\sigma)$ be an OL-system. A symbol a in Σ is called useless if $L(G) \subseteq (\Sigma-\{a\})^*$.

Lemma 5 Let L_1 and L_2 be infinite languages over Σ_1^+ and Σ_2^+, respectively, where $\Sigma_1 \cap \Sigma_2 = \phi$. Let $\min(L_1) > 1$ and $\min(L_2) > 1$. If $L_1 \cup L_2 = t(L(G))$ where $G = (\Gamma,P,\sigma)$ is a POL-system without useless symbols and t is an ℓ-homomorphism, then Γ is the union of two disjoint alphabets Γ_1 and Γ_2, $t(\Gamma_1) = \Sigma_1$, $t(\Gamma_2) = \Sigma_2$, and for each $a \in \Gamma_1$, or $a \in \Gamma_2$ there exists $k_a > 0$ such that if $a \overset{j}{\underset{G}{\Rightarrow}} \alpha$ for $j < k_a$ then $\alpha \in \Gamma_1^+$, or $\alpha \in \Gamma_2^+$, respectively; and if $a \overset{k_a}{\underset{G}{\Rightarrow}} \alpha$, then $\alpha \in \Gamma_2^+$ or $\alpha \in \Gamma_1^+$, respectively.

Proof Let $L_1 \cup L_2 = t(L(G))$ where t is an ℓ-homomorphism and $G = (\Gamma,P,\sigma)$ is a POL-system without useless symbols. Let $\Gamma_1 = t^{-1}(\Sigma_1)$, $\Gamma_2 = t^{-1}(\Sigma_2)$, where t^{-1} is the inverse of homomorphism t. Clearly $\Gamma_1 \cap \Gamma_2 = \phi$, and any string of $L(G)$ is either in Γ_1^+ or in Γ_2^+.

Assume $a \in \Gamma_1$. Since a is not a useless symbol and since $\max(L_1) > 1$ and $\max(L_2) > 1$, there exist $\alpha,\beta \in \Gamma_1^+$, which are not both empty, such that $\alpha a \beta \in L(G)$. It is also clear that for any $k \geq 0$ we cannot derive from a in k steps a string which would contain symbols of both Γ_1 and Γ_2.

Before we complete the proof we need to prove two auxiliary assertions.

Assertion 1. There is no $k \geq 0$ such that $a \overset{k}{\underset{G}{\Rightarrow}} \gamma_1 \in \Gamma_1^+$ and also $a \overset{k}{\underset{G}{\Rightarrow}} \gamma_2$ for $\gamma_2 \in \Gamma_2^+$.

Suppose the negation of assertion 1 is true. Let α_1, β_1 be strings such that $\alpha \overset{k}{\underset{G}{\Rightarrow}} \alpha_1$, $\beta \overset{k}{\underset{G}{\Rightarrow}} \beta_1$. Then $\alpha a \beta \overset{k}{\underset{G}{\Rightarrow}} \alpha_1 \Gamma_1 \beta_1$ and also $\alpha a \beta \overset{k}{\underset{G}{\Rightarrow}} \alpha_1 \gamma_2 \beta_1$. However, α_1, β_1 cannot be both empty and therefore either $\alpha_1 \gamma_1 \beta_1$ or $\alpha_1 \gamma_2 \beta_2$ contains symbols from both Γ_1 and Γ_2 which is a contradiction.

Assertion 2. For every a in Γ_1 there exist α in Γ_2^+ such that $a \underset{G}{\Rightarrow}^+ \alpha$.

Suppose there is a symbol $a \in \Gamma_1$ such that if $a \overset{k}{\underset{G}{\Rightarrow}} \gamma$ for $k \geq 0$ then $\gamma \in \Gamma_1^+$. Since $\alpha a \beta \in L(G)$ there exists $k_1 \geq 0$ such that $\sigma \overset{k_1}{\underset{G}{\Rightarrow}} \alpha a \beta$. Language L_2 is infinite and so for any $j \geq 0$ there exists $\gamma_1 \in L_2$ such that $\sigma \overset{i}{\underset{G}{\Rightarrow}} \gamma'$, $i \geq j$ and $t(\gamma') = \gamma_1$. Let $\delta \in \Gamma_2^+$ be a string which can be derived from σ using s steps, $s \geq k_1$. Hence, $\sigma \overset{s}{\underset{G}{\Rightarrow}} \delta \in \Gamma_2^+$ and also $\sigma \overset{k_1}{\underset{G}{\Rightarrow}} \alpha a \beta \overset{s-k_1}{\underset{G}{\Longrightarrow}} \alpha' \alpha_1 \beta'$ where $\alpha \overset{s-k_1}{\underset{G}{\Longrightarrow}} \alpha'$, $a \overset{s-k_1}{\underset{G}{\Longrightarrow}} \alpha_1$, $\beta \overset{s-k_1}{\underset{G}{\Longrightarrow}} \beta'$, and $\alpha_1 \in \Gamma_1^+$. However, $|\sigma| > 1$ and system G is propagating. Therefore, we would be able to derive from σ in s steps a string containing symbols from δ and from α_1 which is a contradiction.

For any $a \in \Gamma_1$ let $A_a = \{k \geq 1 : a \overset{k}{\underset{G}{\Rightarrow}} \alpha$ for some α in $\Gamma_2^+\}$. From Assertion 2 follows that A_a is nonempty for any $a \in \Gamma_1$. Let $k_a = \min A_a$. It follows from Assertion 1 and from the definition of A_a that k_a has all required properties.

If $a \in \Gamma_2$ then we can repeat the proof above replacing everywhere Γ_2 by Γ_1 and Γ_1 by Γ_2. $\square$

<u>Lemma 6</u> Let L_1, L_2 be essentially deterministic hPOL-languages, $L_1 \subseteq \Sigma_1^+$, $L_2 \subseteq \Sigma_2^+$ and $\Sigma_1^+ \cap \Sigma_2^+ = \phi$. Let $\min(L_1) = \min(L_2) > 1$. If $L_1 \cup L_2$ is in $\mathcal{L}_{hPOL}$ then $L_1 \cup L_2$ is also an essentially deterministic hPOL-language.

<u>Proof</u> Suppose that $L_1 \cup L_2 \in \mathcal{L}_{hPOL}$ but $L_1 \cup L_2$ is not an essentially deterministic hPOL-language. It means that there exist an ℓhomomorphism t and a POL-system $G = (\Gamma, P, \sigma)$ such that G is an essentially nondeterministic system with respect to t and $L_1 \cup L_2 = t(L(G))$.

By definition of an essentially nondeterministic system there exists $a \in \Gamma$ such that:

(i) there are infinitely many words of $L(G)$ containing symbol a;

(ii) there exist $k > 0$, $\gamma_1, \gamma_2 \in \Gamma^+$ such that $a \overset{k}{\Rightarrow} \gamma_1$, $a \overset{k}{\Rightarrow} \gamma_2$ and $t(\gamma_1) \neq t(\gamma_2)$.

Suppose that $t(\gamma_1)$, $t(\gamma_2) \in \Sigma_1^+$. In addition suppose that $t(\sigma) \in L_1$. (In the case $t(\sigma) \in L_2$ we may easily modify G and t to have ℓh-image of starting string in L_1, because $\min(L_1) = \min(L_2)$).

Case 1. Suppose that $t(a) \in \Sigma_1$. By Lemma 5 for every $b \in \Gamma$, $t(b) \in \Sigma_1$, there exists $k_b \geq 0$ such that if $b \overset{j}{\Rightarrow} \alpha$ and $0 \leq j \leq k_b - 1$ then $t(\alpha) \in \Sigma_2^+$ then $t(\alpha) \in \Sigma_2^+$ and if $b \overset{k_b}{\Rightarrow} \alpha$, $\alpha \in \Gamma^+$ then $t(\alpha) \in \Sigma_1^+$. Construct POL-system $G' = (\Gamma', P', \sigma)$ where $\Gamma' = \{b \in \Gamma : t(b) \in \Sigma_1\}$,

$P' = \{b \to \delta : b \in \Gamma', b \to \delta \in P_1, k_b > 1\} \cup \{b \to \delta : b \in \Gamma', k_b = 1, b \to \alpha_1 c \alpha_2 \in P$ for some $\alpha_1, \alpha_2 \in \Gamma^*$ and $c \in \Gamma$, and $b \xRightarrow{k_c+1} \delta$ for some $\delta \in \Gamma^+\}$.

Then $t(L(G')) = t(L(G)) \cap \Sigma_1^+ = L_1$. Since $a \in \Gamma'$ and a satisfies (i) and (ii), G' is an essentially nondeterministic system with respect to t' generating L_1 which is a contradiction to our assumption.

Case 2. Suppose that $t(a) \in \Sigma_2$. Let $A = \{\alpha a\beta : \alpha a\beta \in L(G) \text{ for some } \alpha,\beta \in \Gamma^*\}$. By (i) A is infinite. Since $t(\sigma) \in \Sigma_1^+$ and $t(\alpha a\beta) \in \Sigma_2^+$ for any $\alpha a\beta \in A$, for every string $\alpha a\beta \in A$ there exists a string $\alpha' a_{\alpha\beta}\beta' \in L(G)$ and $j_{\alpha\beta} \geq 1$ which satisfy the following conditions:

(2.1) $\alpha',\beta' \in \Gamma^*$ and $a_{\alpha\beta} \in \Gamma$, $t(\alpha' a_{\alpha\beta}\beta') \in L_1$;

(2.2) $\alpha' a_{\alpha\beta}\beta' \xRightarrow{j_{\alpha\beta}} \alpha a\beta \in A$ and if $\alpha' a_{\alpha\beta}\beta' \xRightarrow{i} \delta$ for $1 \leq i \leq j_{\alpha\beta}$ then $t(\delta) \in L_2$;

(2.3) There exist strings α_1, α_2, β_1, $\beta_2 \in \Gamma^*$ such that $\alpha = \alpha_1\alpha_2$, $\beta = \beta_1\beta_2$ and $\alpha' \xRightarrow{j_{\alpha\beta}} \alpha_1$, $\beta' \xRightarrow{j_{\alpha\beta}} \beta_2$ and $a_{\alpha\beta} \xRightarrow{j_{\alpha\beta}} \alpha_2 a\beta_1$.

For every $b \in \Gamma$ let k_b has the same meaning as in Lemma 5. Clearly, for any $\alpha a\beta \in A$, $j_{\alpha\beta} \leq \max\{k_b : b \in \Gamma\}$. Let $B = \{\alpha' a_{\alpha\beta}\beta' \in L(G) : \alpha' a_{\alpha\beta}\beta'$ satisfies conditions (2.1)-(2.3)$\}$.

Since A is infinite and any string of A can be derived from a string of B in limited number of steps, B must be also infinite. Thus, there exists $a' \in \Gamma$ such that the set $\{\alpha' a'\beta' \in B : \alpha',\beta' \in \Gamma^*\}$ is infinite. There exist $i \geq 1$ and $\delta_1,\delta_2 \in \Gamma^*$ such that $a' \xRightarrow{i} \delta_1 a\delta_2$. Let δ_1', δ_2' be any strings of Γ^* such that $\delta_1 \xRightarrow{k} \delta_1'$, $\delta_2 \xRightarrow{k} \delta_2'$. Then $a' \xRightarrow{i} \delta_1 a\delta_2 \xRightarrow{k} \delta_1'\gamma_1\delta_2'$ and also $a' \xRightarrow{i+k} \delta_1'\gamma_2\delta_2'$. Since $t(\gamma_1) \neq t(\gamma_2)$, we have $t(\delta_1'\gamma_1\delta_2') \neq t(\delta_1'\gamma_2\delta_2')$, and $t(a') \in \Sigma_1$. Thus, if replacing a by a' and γ_1,γ_2 by $\delta_1'\gamma_1\delta_2'$, $\delta_1'\gamma_2\delta_2'$, respectively, in Case 1 of this proof we obtain also a contradiction. $\square$

Lemma 7 The language $L = \{a^{3^n} : n \geq 1\}$ is an essentially deterministic hPOL-language.

Proof Suppose that L is not an essentially deterministic language, i.e. there exist ℓ-homomorphism t and POL-system $G = (\Sigma,P,\sigma)$ which is essentially nondeterministic with respect to t such that $t(L(G)) = L$. It means that there exists symbol $a \in \Sigma$ such that

(i) There are infinitely many words of L(G) where symbol a occurs.

(ii) There exist $k > 0$, $\gamma_1,\gamma_2 \in \Sigma^+$ such that $a \overset{k}{\Rightarrow} \gamma_1$, $a \overset{k}{\Rightarrow} \gamma_2$ and $t(\gamma_1) \neq t(\gamma_2)$.

Let $\alpha_1 a\beta_1 \in L(G)$. There exist $\alpha_2,\beta_2 \in \Sigma^*$ such that $\alpha_1 \overset{k}{\Rightarrow} \alpha_2$ and $\beta_1 \overset{k}{\Rightarrow} \beta_2$. So $\alpha_1 a\beta_1 \overset{k}{\Rightarrow} \alpha_2\gamma_1\beta_2$ and also $\alpha_1 a\beta_1 \overset{k}{\Rightarrow} \alpha_2\gamma_2\beta_2$. Since $\alpha_2\gamma_1\beta_2$ and $\alpha_2\gamma_2\beta_2$ are distinct and both are in L(G), $t(\alpha_2\gamma_1\beta) = a^{3^i}$ and $t(\alpha_2\gamma_2\beta_2) = a^{3^j}$, for some $i,j > 0$, $i \neq j$. We may assume, without loss of generality, that $i > j$. Then $t(\gamma_1) = a^{3^i-3^j}t(\gamma_2)$. It follows from (i) that there exist $\alpha,\beta \in \Sigma^*$ such that $\alpha a\beta \in L(G)$ and $|\alpha a\beta| > 3^i$. Let $\alpha',\beta' \in \Sigma^*$ be strings such that $\alpha \overset{k}{\Rightarrow} \alpha'$ and $\beta \overset{k}{\Rightarrow} \beta'$. Then $\alpha a\beta \overset{k}{\Rightarrow} \alpha'\gamma_2\beta'$. Since G is a propagating OL-system, $|\alpha'\gamma_2\beta'| = 3^m$ for some $m > i$. From the string $\alpha a\beta$ we can derive in k steps also string $\alpha'\gamma_1\beta'$. However, $3^m < |\alpha'\gamma_1\beta'| = 3^m + 3^i - 3^j < 3^{m+1}$, which is a contradiction since $t(\alpha'\gamma_1\beta')$ is in L and thus $|\alpha'\gamma_1\beta'| = 3^s$ for some $s \geq 1$. $\square$

Lemma 8 The language $L = \{b^n c^n d^n : n \geq 1\}$ is an essentially deterministic hPOL-language.

Proof Suppose that L is not an essentially deterministic hPOL-language. It means that there exist POL-system $G = (\Sigma, P, \sigma)$, ℓ-homomorphism t and a symbol $a \in \Sigma$ such that

(i) $L = t(L(G))$.

(ii) Symbol a occurs in infinitely many words of $L(G)$.

(iii) There exist $k > 0$ and $\gamma_1, \gamma_2 \in \Sigma^+$ such that $a \overset{k}{\Rightarrow} \gamma_1$, $a \overset{k}{\Rightarrow} \gamma_2$ and $t(\gamma_1) \neq t(\gamma_2)$.

Let α, β be strings over Σ^* such that $\alpha a \beta \in L(G)$. Let α_1, β_1 be some strings in Σ^* such that $\alpha \overset{k}{\Rightarrow} \alpha_1$, and $\beta \overset{k}{\Rightarrow} \beta_1$. Then $\alpha a \beta \overset{k}{\Rightarrow} \alpha_1 \gamma_1 \beta_1$ and also $\alpha a \beta \overset{k}{\Rightarrow} \alpha_1 \gamma_1 \beta_1$.

Since $t(\gamma_1) \neq t(\gamma_2)$, $\alpha_1 \gamma_1 \beta_1$, $\alpha_1 \delta_2 \beta_1 \in L(G)$, there exist $i, j \geq 1$, $i \neq j$ such that $t(\alpha_1 \gamma_1 \beta_1) = b^i c^i d^i$ and $t(\alpha_1 \gamma_2 \beta_2) = b^j c^j d^j$. We may suppose that $j > i$.

Case 1. Let $t(\gamma_1) \in \{b,c,d\}^+ - b^+ c^+ d^+$. In this case, since $t(\gamma_1) \neq t(\gamma_2)$, $t(\alpha_1 \gamma_1 \beta_1)$ and $t(\alpha_1 \gamma_2 \beta_1)$ cannot be both in L_2.

Case 2. Let $t(\gamma_1) = b^{s_1} c^{s_2} d^{s_3}$ for $s_1, s_2, s_3 > 0$. Then $s_2 = i$, $t(\alpha_1) = b^{i-s_1}$ and $t(\beta_1) = d^{i-s_3}$. Since $t(\alpha_1 \gamma_2 \beta_1) = b^{i-s_1} t(\gamma_2) d^{i-s_3} =$ $= b^j c^j d^j$ thus $t(\gamma_2) = b^{j-i+s_1} c^j d^{j-i+s_3}$. Let α_2, β_2 be strings in Σ^* such that $\alpha_2 a \beta_2 \in L(G)$ and $|\alpha a \beta| > 3j$. By (ii) such strings exist. Let α_3, β_3 be strings which are derived in G from α_2, β_2 respectively, in k steps. Then $\alpha_2 a \beta_2 \underset{G}{\overset{k}{\Rightarrow}} \alpha_3 \gamma_1 \beta_3$ and since G is a propagating system, $|\alpha_3 \gamma_1 \beta_3| > 3j$.

However, $t(\alpha_3\gamma_1\beta_3) = t(\alpha_3)t(\gamma_1)t(\beta_3) = t(\alpha_3)b^{s_1}c^{s_2}d^{s_3}t(\beta_3)$ where $s_2 = i$ and $s_1, s_3 > 0$. Thus $s_2 < |\alpha_3\gamma_1\beta_3| / 3$ which is in a contradiction to $t(\alpha_3\gamma_1\beta_3) \in t(L(G))$. $\square$

<u>Lemma 9</u> The language $L = \{a^{3^i} : i \geq 1\} \cup \{b^i c^i d^i : i \geq 1\}$ is not an hPOL-language.

<u>Proof</u> By Lemma 7 and 9, the languages $L_1 = \{a^{3^i} : i \geq 1\}$ and $L_2 = \{b^i c^i d^i : i \geq 1\}$ are both essentially deterministic hPOL-languages and, clearly, $\min(L_1) = \min(L_2) = 3 > 1$. So by Lemma 6 if L is an hPOL-language then L is an essentially deterministic hPOL-language. Assume that L is an essentially deterministic language. Let $G = (\Sigma, P, \sigma)$ be a POL-system, t be an ℓ-homomorphism, $t(L(G)) = L$. We can suppose without loss of generality that $t(\sigma) \in L_1$ (otherwise we can easily modify G and t such that $t(\sigma) \in L_1$). Let $\alpha \in L(G)$, $t(\alpha) \in L_2$ so that

$$3^i < |\alpha| < 3^{i+1}. \tag{1}$$

Since $t(\sigma) \in L_1$ and $\sigma \Rightarrow^* \alpha$, there exist $n_1 \geq 1$ and $\beta_1 \in \Sigma^+$ such that $\sigma \Rightarrow^* \beta_1 \overset{n_1}{\Rightarrow} \alpha$, $t(\beta_1) \in L_1$ and if $\beta_1 \overset{i}{\Rightarrow} \gamma$ for $0 < i \leq n_1$, $\gamma \in \Sigma^+$ then $t(\gamma) \in L_2$. It follows from Lemma 5 that there exist $\beta_2 \in L(G)$, $n_2 \geq 1$ such that $t(\beta_2) \in L_1$, $\alpha \overset{n_2}{\Rightarrow} \beta_2$ and if $\alpha \overset{i}{\Rightarrow} \gamma$ for $i < n_2$ then $t(\gamma) \in L_2$. Since $\beta_1, \beta_2 \in L_1$ and because of (1), $|\beta| \leq 3^i$ and $|\beta_2| \geq 3^{i+1}$. Also by Lemma 5 there exists constant c, independent on i, such that

$$n_1 + n_2 < c. \tag{2}$$

It follows from Lemma 4 that there exist $n \geq 1$, PDOL-systems $G_1, G_2, \ldots, G_n$ and ℓ-homomorphismus t' such that $L(G) = \bigcup_{j=1}^{n} t(t'(L(G_i))$ and if $\gamma \overset{s}{\underset{G_j}{\Rightarrow}} \delta$ for some $1 \le j \le n$, $s \ge 0$ then $t'(\gamma) \overset{s}{\underset{G}{\Rightarrow}} t'(\delta)$. Let composition $t \circ t'$ be denoted by t''. Let for any $i \ge 1$, $A_i = \{(\beta_1,\beta_2): t''(\beta_1), t''(\beta_2) \in L_1$, $|\beta_1| \le 3^i$, $|\beta_2| \ge 3^{i+1}$, $\beta_1 \overset{k}{\underset{G_j}{\Rightarrow}} \beta_2$ for some j, $1 \le j \le m$, $k \ge 1$, and if $\beta_1 \overset{m}{\underset{G_j}{\Rightarrow}} \gamma$ for $0 < m \le j$ then $t''(\gamma) \in L_2\}$.

It follows from the discussion above that for any string $\alpha \in L(G)$, $t'(\alpha) \in L_2$, $3^i < |\alpha| < 3^{i+1}$, there exists $(\beta_1,\beta_2) \in A_i$ such that $\beta_1 \Rightarrow^* \alpha \Rightarrow^* \beta_2$. Clearly, because every G_j is a deterministic propagating system for every $1 \le j \le n$, $L(G_j)$ can contain at most two strings δ_1,δ_2 such tha $(\delta_1,\delta_2) \in A_i$ and thus $||A_i|| \le n$.

It follows from (2) that for any $(\beta_1,\beta_2) \in A_i$ there are at most c-2 strings whose ℓh-images are in L_2, whose length are between 3^i and 3^{i+1} and which can be derived from β_1. Therefore, we have $||\bigcup_{j=3^{i+1}}^{3^{i+1}-1} L_2^{(j)}|| \le n \cdot c$ for any $i \ge 1$. However, it follows from the definition of language L_2 that $||\bigcup_{j=3^{i+1}}^{3^{i+1}-1} L_2^{(j)}|| = 3^i - 3^{i-1}$ for $i \ge 1$ which is a contradication. $\square$

<u>Theorem 9</u> $\mathcal{L}_{hPD0L} \subsetneq \mathcal{L}_{hPDF0L}$.

<u>Proof</u> Languages L_1, L_2 from the proof of Lemma 9 are both in $\mathcal{L}_{hPD0L}$ but their union is not in $\mathcal{L}_{hPD0L}$ as shown in Lemma 9. Clearly, $L_1 \cup L_2$ is in $\mathcal{L}_{hPDF0L}$. $\square$

Theorem 10 $\mathcal{L}_{hPOL} \subsetneq \mathcal{L}_{hPFOL}$.

Proof Directly by Lemma 9 since $L_1 \cup L_2$ is in $\mathcal{L}_{hPFOL}$. □

Theorem 11 $\mathcal{L}_{hPDFOL}$ and $\mathcal{L}_{hPOL}$ are incomparable.

Proof The language a^+b^+ is in $\mathcal{L}_{hPOL}$ but not in $\mathcal{L}_{hPDFOL}$ as shown in the proof of Lemma 2. The language $L_1 \cup L_2$ from the proof of Lemma 9 is in $\mathcal{L}_{hPDFOL}$ but not in $\mathcal{L}_{hPOL}$. □

The inclusion results on the considered families of languages are summarized in Figure 1. The meaning of the graph is the following. If two nodes labeled $\mathcal{L}_A$, $\mathcal{L}_B$ are connected by an edge (double edge), the node $\mathcal{L}_B$ being below the node $\mathcal{L}_A$, then $\mathcal{L}_B \subseteq \mathcal{L}_A$, ($\mathcal{L}_B \subsetneq \mathcal{L}_A$). If two nodes labeled $\mathcal{L}_A$, $\mathcal{L}_B$ are connected by a broken edge, then $\mathcal{L}_B$, $\mathcal{L}_A$ are incomparable. If two nodes are not connected by any edge or ascending unbroken path then their relation is open.

5. Closure properties of hOL-languages Some closure properties of $\mathcal{L}_{hOL}$ are given in Table 1. It is easy to show, using similar techniques as in the proof of Theorems 1 and 3 that $\mathcal{L}_{hOL}$ is closed under concatenation, operation star and ε-free homomorphism. By Theorem 1, $\mathcal{L}_{hOL}$ is closed under union. It follows from [5] that $\mathcal{L}_{hOL}$ is not closed under inverse homomorphism.

	$\cup$	$\cdot$	$*$	ε-free h	h^{-1}
$\mathcal{L}_{hOL}$	YES	YES	YES	YES	No

Table 1.

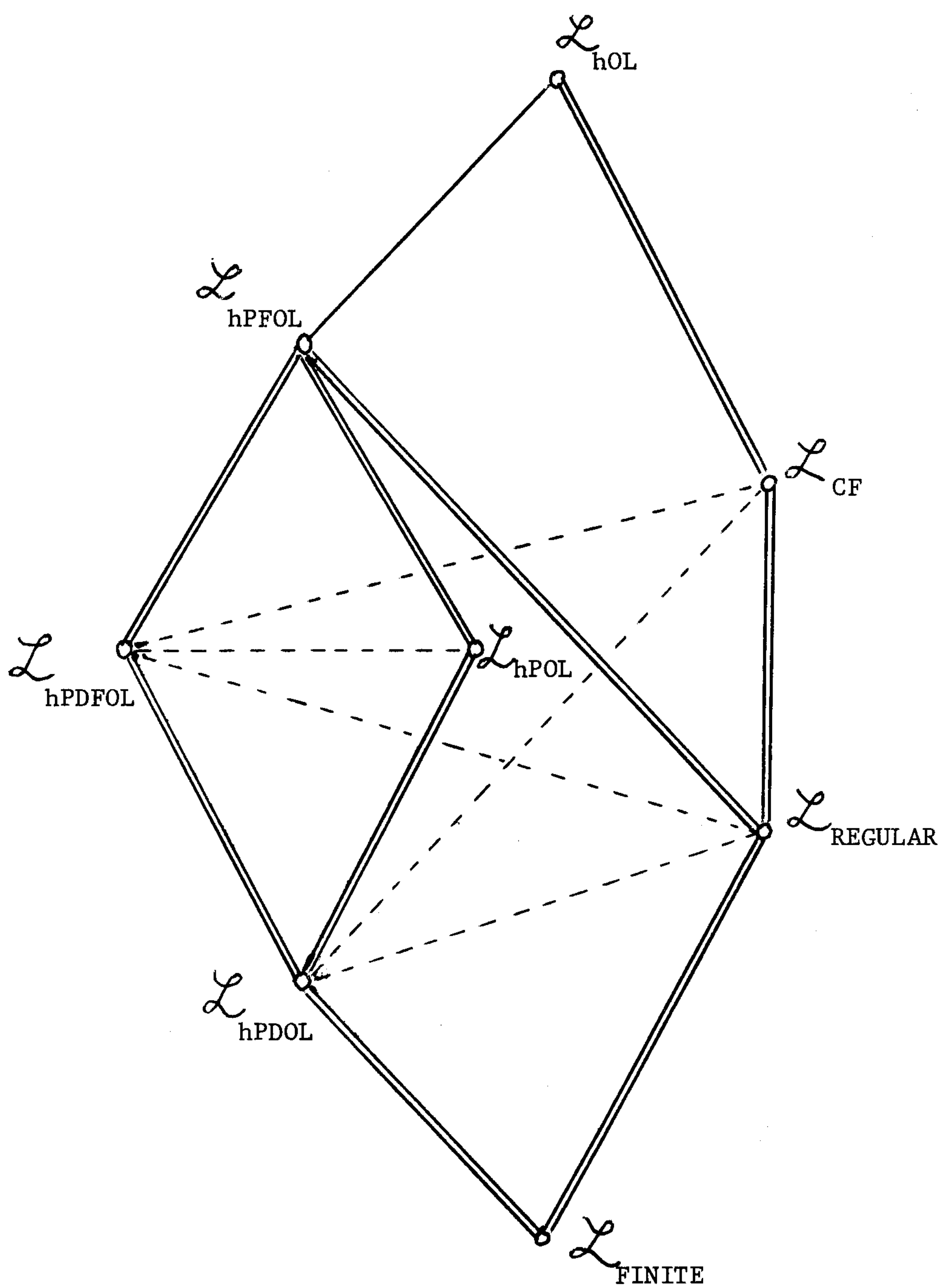
$\mathcal{L}_{hOL}$
$\mathcal{L}_{hPFOL}$
$\mathcal{L}_{CF}$
$\mathcal{L}_{hPDFOL}$
$\mathcal{L}_{hPOL}$
$\mathcal{L}_{REGULAR}$
$\mathcal{L}_{hPDOL}$
$\mathcal{L}_{FINITE}$

Figure 1.

References

[1] Aho, A.V., Indexed Grammars - An Extension of Context-free Grammars, J. Assoc. Comput. Mach., 15(1968), 647-671.

[2] Culik, K. II and Opatrny, J., Macro OL-systems, Research Report CS-73-06, University of Waterloo, Waterloo, Ontario, Canada, 1973.

[3] Culik, K. II, On Some Families of Languages related to Developmental Systems, Int. J. of Comp. Math., to appear. Also Research Report CS-73-12, University of Waterloo, Waterloo, Ontario, Canada, 1973.

[4] Ginsburg, S. and Greibach, S., Abstract Families of Languages, Memoirs Amer. Math. Soc., 87 (1969), 1-12.

[5] Herman, G.T., Closure Properties of Some Families of Languages Associated with Biological Systems, Int. J. of Comp. Math., to appear. Also in Proceedings of the Fifth Annual Princeton Conference on Information Sciences, Princeton, 1971.

[6] Hopcroft, J.E. and Ullman, J.D., Formal Languages and their Relation to Automata, Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

[7] Lindenmayer, A., Mathematical Models for Cellular Interactions in Development, J. Theoret. Biol., 18(1968), 280-315.

[8] Lindenmayer, A., Developmental Systems without Cellular Interactions, their Languages and Grammars, J. Theoret. Biol., 30(1971), 455-484.

[9] Rozenberg, G. and Doucet, P.G., On OL-languages, Information and Control, 19(1971), 303-318.

[10] Rozenberg, G., Extension of Tabled OL-systems and Languages. Research Report 53-73, Computer Science Department, SUNY at Buffalo, 1972.

[11] Salomaa, A., Formal Languages, Academic Press, New York, 1973.